UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/521,732	01/19/2005	Li-Qun Xu	36-1883	6894

23117 7590 11/05/2008
NIXON & VANDERHYE, PC
901 NORTH GLEBE ROAD, 11TH FLOOR
ARLINGTON, VA 22203

EXAMINER	
OLUDE AFOLABI, OLATOYOSI	
ART UNIT	PAPER NUMBER
2455	
MAIL DATE	DELIVERY MODE
11/05/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/521,732	XU ET AL.	
	Examiner	Art Unit	
	OLA OLUDE AFOLABI	2455	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *30 October 2006*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *11-18* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *11-18* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *10/30/06* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This communication is considered fully responsive to the Amendment filed on 9/24/2008 for the patent application 10/521,732. Claims 11-18 have been examined and remain pending.

Continued Examination Under 37 CFR 1.114

2. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on September 30, 2008 has been entered.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claims 11-18 are rejected under 35 U.S.C. 103(a) as being unpatentable over C. C. Chibelushi et al. (*A Review of Speech-Based Bimodal Recognition*) (**Chibelushi** hereafter) and **Haykin** (*Neural Networks: A Comprehensive Foundation, Chapter 9*)

Regarding Claim 11, and 15:

Chibelushi discloses:

- feature extraction means (e.g. *Mouth-window methods,* col. 1: par. 2a, pp. 26, Chibelushi) for extracting by a feature extraction module (e.g. MFCC feature extraction, Fig. 2) a plurality of sets of characteristic visual feature vectors and a plurality of sets of characteristic audio feature vectors from respective video and audio portions of a training set comprising a plurality of video sequences belonging to a predetermined class (**Note:** This is a variation of the learning scheme of neurons, which makes the Kohonen network into a classification network called Learning Vector Quantization (LVQ), the modification involves changing the training scheme, which requires a collection of training examples each assigned to one of a set of known classes, see Tsoukalas pp. 314);
- feature combining means (e.g. sensor fusion, Low-level fusion can occur at the data level or feature level. Intermediate-level and high-level fusion involves the combination of recognition scores or labels produced as intermediate or final output of classifiers, col. 1 Section B pp. 28) for combining by a feature binder the plurality of sets of characteristic visual and audio feature vectors into a respective plurality of N-dimensional feature vectors corresponding to the predetermined class (e.g. audio-visual fusion can also occur at a level between feature and decision levels, fig. 1), said combining comprising normalizing and concatenating each of the visual

feature vectors with corresponding audio feature vectors (e.g. Some similarity measures are tightly coupled to particular feature types. For speaker verification or open-set identification, a normalization of similarity scores may be necessitated by speech variability [63], [101]. Examples of common similarity measures are: the Euclidean distance (often inverse-variance weighted, or reduced to a city-block distance, col. 2, section C, pp. 27- col. 1: 1-8, pp. 28); (**Note:** This is a variation of the learning scheme of neurons, which makes the Kohonen network into a classification network called Learning Vector Quantization (LVQ), the modification involves changing the training scheme, which requires a collection of training examples each assigned to one of a set of known classes, see Tsoukalas pp. 314);

- analysing by a feature learning module the pluralities of N-dimensional feature vectors (sometimes called n-dimensional weight vector, see Tsoukalas pp. 309) using principal component analysis (e.g. Some high-level features aim at reducing dimensionality through a transformation (e.g. transforms are based on *principal component analysis (PCA)*, statistical discriminant analysis optimizing the F-ratio such as *linear discriminant analysis (LDA)* [1], and *integrated mel-scale representation with LDA (IMELDA)*) that produces statistically orthogonal features and packs most of the variance into few features, col. 1, pp. 25) or kernel discriminant analysis to generate a set of M basis vectors (sometimes referred to as m-dimensional input vector, see Tsoukalas pp. 309), each being of N-dimensions (e.g. applied to static spectral information, possibly

combined with dynamic spectral information, output by a mel-scale filter bank. Composite features are sometimes generated by a simple concatenation of different types of features, col. 2 pp. 25)

Chibelushi fails to particularly call for:

- a plurality of sets of characteristic visual feature vectors and a plurality of sets of characteristic audio feature vectors from respective video and audio portions of a training set comprising a plurality of video sequences
- wherein M << N, and using the set of M basis vectors, mapping each N-dimensional feature vector into a respective M-dimensional feature vector
- using the M-dimensional feature vectors (**Note:** feature vectors are corresponding inputs as the components of an m-dimensional input vector, see Tsoukalas pp. 309) thus obtained as the basis for or as input to train a class model of the predetermined class
- storing the class model for use in classifying input data that matches the predetermined class

Haykin teaches:

- a plurality of sets of characteristic visual feature vectors and a plurality of sets of characteristic audio feature vectors from respective video and audio portions of a training set (e.g. input pattern presented to the network, (pp. 443-483 especially pp.447, Haykin) comprising a plurality of video sequences (e.g. data from input space, fig. 9.4 pp. 455, Haykin).
- wherein M << N, and using the set of M basis vectors, mapping each N-dimensional feature vector into a respective M-dimensional feature vector (**Note:** continuous input space

is mapped to discrete output space through the feature map, fig. 9.4 pp. 455, Haykin);

- using the M-dimensional feature vectors thus obtained as the basis for or as input to train a class model of the predetermined class (e.g. input pattern presented to the network, (pp. 443-483 especially pp.447, Haykin).
- storing the class model for use in classifying input data that matches the predetermined class (**Note:** aim of the SOM algorithm is to store a large set of input vectors by finding a smaller set of prototypes, so as to provide a good approximation to the original input space, see par. 2, pp. 455, Haykin).

Rationale:

Thus, it would have been recognized by one of ordinary skill in the art at the time of the invention to modify the teachings of **Chibelushi** for generating class models from video sequences having one of a plurality of predetermined classes with the teachings of **Haykin** for the benefit of reducing dimensionality or compressing data and also to store a large set of input vectors by finding a smaller set of prototypes, so as to provide a good approximation to the original input space.

Regarding Claims 12 and 16, The computer-implemented method as claimed in claim 11, wherein the M basis vectors are the M most discriminating basis vectors that maximize between-class variance and minimize within-class variance (e.g. conditions for the minimization of the expected distortion which is given the input x, choose the code c = c(x) to minimize the squared error distortion ||x - x'||2, pp. 456, Haykin).

(**Note**: In applying Haykin's theory to audio-video skims, the objective function that is to be minimized has constraints (audio/video duration constraints, visual syntax, synchronous multimedia constraints) that are constructed with the aim of maximizing the speech information content and the overall coherence of the video, see col. 1, par. 3, pp. 2, Sundaram).

Regarding Claims 13 and 17, The computer-implemented method as claimed in claim 11 wherein each video sequence has a non-linear feature distribution (e.g. video sequence is a feature map, (Fig. 9.7b pp. 462), and the property 4 of a feature map is feature selection, see pp. 461, Haykin).

Regarding Claims 14 and 18, The computer-implemented method as claimed in claim 12 wherein each video sequence has a non-linear feature distribution (e.g. video sequence is a feature map, (Fig. 9.7b pp. 462), and the property 4 of a feature map is feature selection i.e. video sequence is data from an input space with a non-linear distribution, pp. 461, Haykin) .

Conclusion / Correspondence Information

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Ola Olude-Afolabi whose telephone number is (571)270-5639. The examiner can normally be reached on Monday-Thursday 9:00 - 5:00.

As detailed in MPEP 502.03, communications via Internet e-mail are at the discretion of the applicant. Without a written authorization by applicant in place, the USPTO will not respond via Internet e-mail to any Internet correspondence which contains information subject to the confidentiality requirement as set forth in 35 U.S.C. 122. A paper copy of such correspondence will be placed in the appropriate patent application. The following is a sample authorization form which may be used by applicant:

"Recognizing that Internet communications are not secure, I hereby authorize the USPTO to communicate with me concerning any subject matter of this application by electronic mail. I understand that a copy of these communications will be made of record in the application file."

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Vincent can be reached on 571-272-3080. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system,

see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Ola Olude-Afolabi,
Examiner,
Art Unit 2455

/O. O./
Examiner,
Art Unit 2455
/David R Vincent/
Supervisory Patent Examiner, Art Unit 2129